



Overview of NCBI Resources

Selected databases, tools, help documents and educational materials

<https://www.ncbi.nlm.nih.gov/>

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

The National Center for Biotechnology Information (NCBI) provides many databases and analysis tools to the research community, most of which are readily accessible through the NCBI homepage at www.ncbi.nlm.nih.gov. Every year, NCBI publishes an update in the special database issue of *Nucleic Acids Research* to provide an updated description of databases, services, and tools available from NCBI. The full-text update for 2020 is available online at: www.ncbi.nlm.nih.gov/pmc/articles/PMC6943063/



Selected NCBI Resources

The table below groups and describes key NCBI databases. Append the text in the URL extension column to the NCBI domain web address (www.ncbi.nlm.nih.gov) to navigate to that particular database.

Literature	URL ext	Description
PubMed	/pubmed	Biomedical literature citations and abstracts
PubMed Central	/pmc	Free, full-text articles in the life sciences
Medical Subject Headings	/mesh	Controlled vocabulary for indexing PubMed and other databases
Books	/books	Biomedical and scientific texts and reports, plus NCBI documentation
NLM Catalog	/nlmcatalog	A list of available NLM resources including a searchable set of journals indexed in PubMed
Clinical Genetic resources		
ClinVar	/clinvar	A public archive of reports of evidence-supported relationships between human variation and phenotype
NIH Genetic Testing Registry (GTR)	/gtr	A database of molecular, biochemical and cytogenetic clinical and research genetic test information voluntarily submitted by laboratories worldwide.
MedGen	/medgen	Clinical genetics data aggregated from authoritative sources with actionable information
Medical Genetics Summary	/books/NBK61999	A collection of summaries describing the known impact specific sequence variations have on health
GeneReviews®	/books/NBK1116	An international point-of-care resource for busy clinicians with actionable information
Projects & Studies		
Taxonomy	/taxonomy	Classification of organisms represented in the NCBI molecular databases
BioProject	/bioproject	Experimental projects organized by taxonomy and study
dbGaP	/gap	Large-scale genotype and phenotype association studies
GEO DataSets	/gds	Functional genomics studies with expression data
PubChem BioAssay	/pcassay	Bioactivity screening studies for selective chemical substances
Nucleotide Sequences		
Nucleotide	/nuccore	Nucleotide sequence records, including Reference Sequences
PopSet	/popset	Sets of sequences from population, phylogenetic, ecosystem, and mutation studies
Sequence Read Archive	/sra	High-throughput sequencing data from next-generation sequencing platforms
Genome & Variation		
Genome	/genome	Records on organisms integrated with links to sequence, genetic and physical maps
Assembly	/assembly	Records describing specific versions of genome assemblies with links to sequences
dbSNP	/snp	Short human genomic variations database containing SNPs and short insertions/deletions
dbVar	/dbvar	Large-scale human genomic variation database containing copy-number variations, as well as insertions and deletions longer than 50 bases



Selected NCBI Databases (cont.)

Genes & Proteins	URL ext	Description
Gene	/gene	Summary pages of gene-centered information
Protein	/protein	Translations of nucleotide sequence records and sequences from PDB, SWISS-PROT, PIR, PRF, and protein sequences from patent granting agencies
Identical Protein Groups	/ipg	Consolidated protein records based on identical sequences
Conserved Domains	/cdd	Catalog of curated protein domains
Structure	/structure	Experimental 3D structure data from the Protein DataBank (PDB)
BioSystems	/biosystems	Biological pathways with links to relevant genes, proteins, and chemicals
Chemicals		
PubChem Compound	/pccompound	Information on chemical compounds

Searching the Databases

You can query NCBI databases with different types of input queries by using different interfaces.

- **Querying with text:** The Entrez system supports keyword searching and integrates databases containing scientific literature, taxonomy, sequence, genome, expression and structure data into a tightly interlinked system. Each database has a discrete set of indexed Fields, which are available for more effective and specific searching. Popular terms in the most commonly used indexed Fields are selectable on specialized “Limits” pages or filter list on the left of the search result. Complex queries can be constructed using functions in the “Advanced” page.
- **Querying with sequence:** The Basic Local Alignment Search Tool (BLAST®) is a local sequence similarity search program that compares nucleotide or protein sequence queries to those in sequence databases. It assigns statistical significance to help decipher the biological importance of the alignments. The BLAST algorithm can be used through a web interface, at blast.ncbi.nlm.nih.gov, and is also available as a package of stand-alone tools available for download from ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/.
- **Querying with macromolecular structure:** Vector Alignment Search Tool (VAST) is a structure similarity search service for macromolecules. VAST compares the three-dimensional position of protein secondary structure elements to those of other structures in the Structure database and allows for viewing superimpositions and alignments using the molecular visualization program Cn3D. The service is available online through this web page: www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml.
- **Querying with chemical structure:** The PubChem Structure Search is a structure similarity search tool for small chemical compounds. It searches for similar chemical structures in the PubChem Compound database based on substructure binary fingerprints using the Tanimoto scoring system. The service is available online at pubchem.ncbi.nlm.nih.gov/search/search.cgi.

You can also access these searches through **API**. Refer to the following documents for more information:

Entrez API www.ncbi.nlm.nih.gov/books/NBK25501/, Datasets API api.ncbi.nlm.nih.gov/datasets/docs/datasets-api/, Variation Service api.ncbi.nlm.nih.gov/variation/v0/ and PubChem pubchemdocs.ncbi.nlm.nih.gov/programmatic-access

Personalized NCBI Experience with My NCBI

After registering for a free account (www.ncbi.nlm.nih.gov/account/register/) and logging in, you can use My NCBI to set display preferences for NCBI Web services. By logging into a My NCBI account, you can use other customizable services, such as:

- 1) saving searches and setting automatic email alerts for newly released content,
- 2) storing and maintaining collections of records, as well as sharing Collections with other My NCBI users,
- 3) setting custom filters for more relevant result sets,
- 4) displaying other relevant resources through links from LinkOut providers,
- 5) creating a Biosketch and MyBibliography for grant application and progress purposes,
- 6) saving BLAST search strategies for future references, and
- 7) checking the status of BLAST searches and accessing their results.

See My NCBI help documentation for more information (www.ncbi.nlm.nih.gov/sites/books/NBK3842/). A collection of usage examples demonstrating different features of MyNCBI is also available (ftp.ncbi.nlm.nih.gov/pub/factsheets/Booklet_Teaching_MyNCBI.pdf).

Data Submission to NCBI


Data submission is accepted by various NCBI databases. The NCBI Submission Portal (submit.ncbi.nlm.nih.gov/) lists databases accepting submission and provides links to relevant tools and online help.

Education Resources, Help Manuals and Glossaries

The Learn Page: This Learn page (shown to the right) is the portal to a wide variety of educational and help resources, such as How-To Guides, with quick directions for specific tasks; Problem Sets with detailed examples for exploration; Tutorials with written directions for using a tool or resource; and links to Video Tutorials from the NCBI Channel on YouTube containing demonstrations of how to use a tool or resource (A). From this page, links to additional help documentation, training materials, and programs are also available (B). These include Resource Fact Sheets, the NCBI Help Manual, Curriculum Materials, Webinars & Workshops, plus Resource-Specific Help and FAQs. Courses offered by NCBI, including the monthly webinars open to the public and archival materials, are listed under the “Webinars & Courses” (C).


Learn

NCBI creates a variety of educational products including courses, workshops, webinars, training materials and documentation. NCBI educational events are free and open to everyone. All NCBI educational materials are available for anyone to re-use and distribute.




Webinars & Courses

In-person courses, live webinars and webinar recordings




Conferences & Presentations

Booth exhibits and workshops at scientific conferences




Tutorials

Tutorials: Training materials in HTML, PDF and video formats



Documentation

Online manuals, handbooks, fact sheets and FAQs



News, Blog & Social Media

Keep up with the latest NCBI news and follow NCBI on social media sites, including FaceBook, Twitter, Google+, LinkedIn and the NCBI Insights blog.

Follow Us

[f](#)
[t](#)
[g+](#)
[in](#)
[You Tube](#)
[RSS](#)
[Email](#)

NCBI News & Blog

Updated and improved collection of RefSeq representative genome assemblies now available
21 Aug 2020

We have updated the collection of representative genome assemblies for Bacteria and Archaea. As announced in April, this set is now recalculated three times a year. We selected a total of 11,727

INSDC Statement on SARS-CoV-2 sequence data sharing during COVID-19
17 Aug 2020

The National Library of Medicine and its partners in the International Nucleotide Database Collaboration (INSDC) have joined together to issue a statement encouraging the scientific

New annotations in RefSeq: white-tufted-ear marmoset, ruddy duck, and more
12 Aug 2020

In June and July, the NCBI Eukaryotic Genome Annotation Pipeline released new annotations in RefSeq for the following organisms: *Acipenser ruthenus* (sterlet) *Anguilla anguilla* (European eel)

[More...](#)

Help Documents: The NCBI Help Manual, available on the NCBI Bookshelf (www.ncbi.nlm.nih.gov/books/NBK3831/), is a set of online books dedicated to providing information and tutorials for a wide range of NCBI resources. Each chapter is devoted to a single service. After a brief overview, there are sections describing how the resource works, detailed directions on how to use the resource, as well as quick tips and tricks. These chapters are often linked directly from the resource homepages and answer most frequently asked questions.

NCBI Handbook: The NCBI Handbook, located on the NCBI Bookshelf (www.ncbi.nlm.nih.gov/books/NBK21101/), is an in-depth guide to NCBI resources for audiences from novice to aspiring power users. All articles can be searched online and downloaded in PDF format. Each chapter is devoted to a specific service. After a brief overview, there is an account of how the resource works, including topics such as how data are included in a database, database design, query processing, and how the different resources relate to each other. Topics include the Entrez Search and Retrieval System, BLAST Sequence Analysis Tool, Processing of Sequence Data at NCBI, Genome Assembly and Annotation at NCBI, and the Reference Sequence (RefSeq) Project. There are also chapters on GenBank, the Taxonomy Project, as well as databases such as Gene, PubMed, PubMed Central, Structure, SNP, and GEO.

NCBI Glossary: Several glossaries are available from NCBI covering the vocabulary of biology and bioinformatics. Representative ones are the NCBI Handbook Glossary (www.ncbi.nlm.nih.gov/books/NBK21106/) and the BLAST Glossary (www.ncbi.nlm.nih.gov/books/NBK62051/).

NCBI Factsheets (short tutorials) & Flyers: The resource factsheets, such as this entry, are archived on the NCBI ftp site (<ftp://ncbi.nlm.nih.gov/pub/factsheets/>). Entries for specific NCBI resources are effectively short tutorials with introduction to the resource as well as instruction and use cases. The readme at <ftp://ncbi.nlm.nih.gov/pub/factsheets/README.html> (shortened to bit.ly/ncbi_factsheets) lists available entries and groups them by content into the following categories: Overview, Factsheets, How-To, Booklet, and Flyers.

NCBI Youtube Channel: This channel (www.youtube.com/user/NCBINLM/) provides access to a broad collection of videos produced by NCBI. These videos covers a wide variety of topics in different formats, from short introduction on features found in NCBI tools, to webinars and workshop presentations. Related entries are also grouped through topic-centric playlists and several such playlists are available.

NCBI Announcements and Social Media

Blog: The NCBI Insights blog (ncbiinsights.ncbi.nlm.nih.gov) provides insights into NCBI resources and the science behind them through briefs, short technical tips, as well as featured articles.

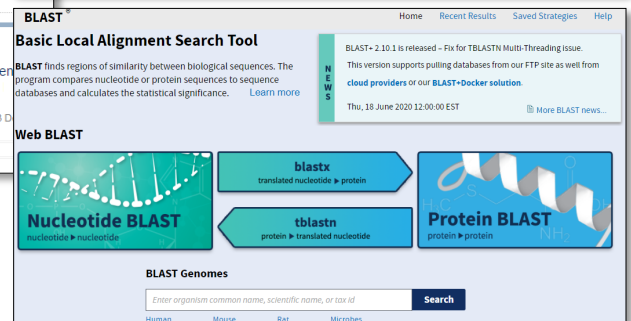
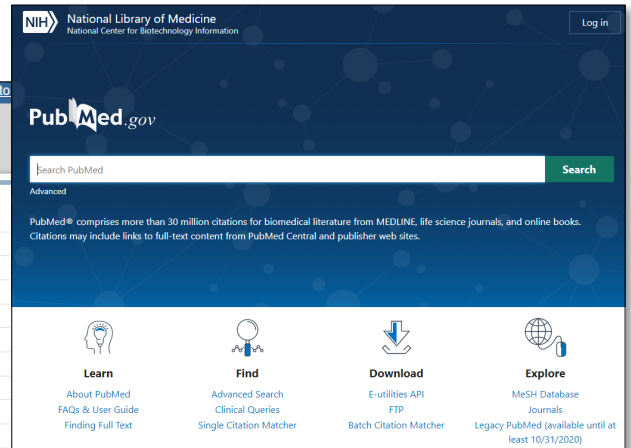
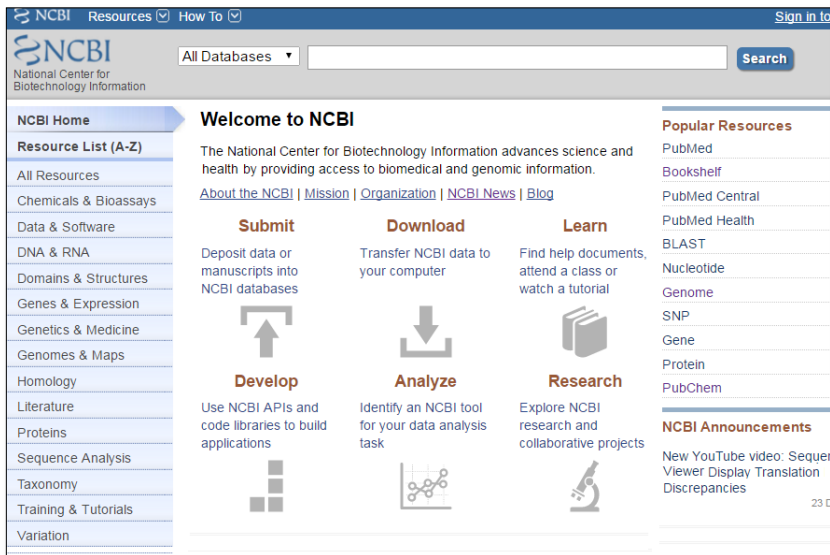
Twitter: The NCBI Twitter Feed provides NCBI announcements, updates, and important calendar events. Follow it at twitter.com/NCBI.

Facebook: The NCBI Facebook page (facebook.com/ncbi.nlm) displays NCBI announcements and provides the opportunity to participate in community discussions.

LinkedIn: The NCBI LinkedIn page (www.linkedin.com/company/national-center-for-biotechnology-information-ncbi/) displays NCBI announcements and recent updates and links to other related pages from NCBI.

RSS Feeds: This page (www.ncbi.nlm.nih.gov/feed/rss.cgi) lists the general as well as resource-specific rss feed.

ListServ: Email ListServes (www.ncbi.nlm.nih.gov/mailman/listinfo/) provide announcements and updates for various NCBI resources. Subscribe to get automatic email notification on important update.



Technical Assistance and More...

NCBI provides extensive technical support to its user community. Email is the most common and often the best method for obtaining technical support. Support is available for help with data submissions, advice on data access, use of specific analytical tools, as well as user feedback and bug reports. Email aliases to handle questions on specialized resources exist. The table below lists the contact information for these specialized resources.

NCBI General User Support

BLAST Help

info@ncbi.nlm.nih.gov

GenBank Submissions/Updates

blast-help@ncbi.nlm.nih.gov

High-throughput Next Generation Sequencing Data

gb-admin@ncbi.nlm.nih.gov

Clinical Data and Authorized Access

sra@ncbi.nlm.nih.gov

Genetic Testing Registry

dbgap-help@ncbi.nlm.nih.gov

Short Genetic Variation (SNP) Data

gtr@ncbi.nlm.nih.gov

Clinical Variation Data

snp-admin@ncbi.nlm.nih.gov

Gene Expression Data

clinvar@ncbi.nlm.nih.gov

PubChem Help

geo@ncbi.nlm.nih.gov

PubMed Central (PMC)

pubchem-help@ncbi.nlm.nih.gov

NIHMS/Open Access Publications

pubmedcentral@ncbi.nlm.nih.gov

NLM General User Support

[nihms-help@ncbi.nlm.nih.gov](mailto:.nihms-help@ncbi.nlm.nih.gov)

custserv@nlm.nih.gov